Engineer

ビジネスインテリジェンス

事業のデータを意思決定に役立つレポートへ変えること

どのようなものか

ビジネスインテリジェンスは、通常BIと略し、組織自身の記録を、意思決定者が行動に使えるレポート、ダッシュボード、分析へ変える仕事です。英語で形容詞として使う場合は、business-intelligence softwareのように、ハイフンを入れることがよくあります。

この語は、ソフトウェアより古いものです。1958年10月、IBM Journal of Research and Developmentは、H. P. Luhnの論文「A Business Intelligence System」を掲載しました。Luhnが説明したのは、組織が受け取る文書や自ら作る文書を取り込み、各文書を自動で要約し、使われる言葉のパターンで特徴を表し、関心のプロファイルが合う「アクションポイント」へ送る機械システムです。アクションポイントは、一人の人でも、グループ、部署、事業部全体でも構いません。businessには、科学や商業から法律、行政、防衛まで、あらゆる組織的な活動を含む、意図して広い意味を与えました。intelligenceは、目の前の事実の関係を捉え、目標へ向けて行動の方向を決められるだけの理解としました。手法を列挙した後、全体を「これをビジネスインテリジェンスシステムと呼ぶ」と名付けています。

Luhnが扱った材料は、文章でした。今日のBIは、主にテーブル内の数値を扱い、Kimball Groupによるデータウェアハウス・BIシステムの構成説明に、典型的な形が示されています。バックルームでは、元のシステムからデータを抽出し、整備し、整合させ、読み込みます。フロントルームでは、業務プロセス別に整理し、最も細かい粒度で保持したデータを提示し、速く使えるよう集計データを加えます。その上に、BIアプリケーションを置きます。個別のクエリ、定型レポート、ダッシュボード、スコアカード、さらに重い分析やデータマイニングのアプリケーションです。

二つの意味を結び付けるのは、Luhnの問いです。誰がこれを知る必要があり、その人が実際に読む形で、どう届けるかということです。

概念を判断に生かす

「ビジネスインテリジェンスとは？」を実際の判断に生かしましょう。その概念を、開発または購入するシステムに結び付けて考えるお手伝いをします。 意思決定を左右する少数の指標を示すダッシュボードやレポートを、依頼しなくても更新される形で提供します。

サービスを見る お問い合わせ →

用語の背後にある前提を検証する

解説とリンク先の実績を参考に、プロジェクトについての質問を整理してください。ツールを選ぶ前に、前提を一緒に検討できます。 読み手の反応に基づくブランド発信 — 読み手の反応に基づくブランド発信

読み手の反応に基づくブランド発信 お問い合わせ →

なぜ重要なのか

どの組織も、すでに数値で動いています。問題は、その数値が一致しているかです。営業が契約の締結日に注文を計上し、経理が請求書の発行日に計上すると、二つの部署が同じ会議へ異なる売上高を持ち込みます。会議は、判断の代わりに、その差の説明に費やされます。BIシステムは、指標ごとに一つの定義を定め、一か所で計算し、尋ねるすべての人に同じ数値を見せることで、費用に見合う価値を生みます。

Luhnが1958年に示した診断は、今の共有レポートフォルダーの大半にも当てはまります。念のために広く情報を送る組織を観察し、受け手が紙に埋もれることで、結局は何も伝わらなくなりやすいと指摘しました。四十のグラフを載せたダッシュボードも、画面上で同じことをします。対策は、選別でした。各アクションポイントには、まず、選んだ文書の番号、名前、著者を載せた短い一覧を届けます。受け手は、その番号をダイヤルして要約を取得し、欲しいものを全文で受け取れました。全文を受け取るたびに関心のプロファイルを更新するため、使うほど選別が改善しました。今のよいレポートも、同じ考え方で動きます。人と、その人が行う判断から始め、その判断に必要な少数の数値を示します。

実務上の理由もあります。アプリケーションのデータベースは、多数の小さな読み書き向けに設計されています。一年分の注文を読む問いを同じ場所で実行すると、その書き込みと資源を取り合います。これが、BIのデータを、定期的に読み込む別の保存先に置くことが多い理由の一つです。Kimballの説明のように、そこで最も細かい粒度のデータを保持するかどうかが、来年、元のシステムへ戻らずに答えられる問いの範囲を決めます。

実際の活用例

典型的なBIの問いを考えます。2026年第3四半期の売上高を、地域別・商品別に示し、各地域の小計と全体の総計を加えるものです。販売一件ごとに一行があり、 `sold_on` 、 `region` 、 `product` 、 `amount` を持つ `sales` テーブルがあれば、

標準SQLの一つのクエリで答えられます。

```
SELECT region,
       product,
       sum(amount)               AS revenue,
       GROUPING(region, product) AS level
FROM   sales
WHERE  sold_on >= DATE '2026-07-01'
AND    sold_on <  DATE '2026-10-01'
GROUP  BY ROLLUP (region, product)
ORDER  BY GROUPING(region), region, GROUPING(product), product;
```

少数のサンプル行に対して実行すると、次の結果になります。

```
 region | product | revenue | level
--------+---------+---------+-------
 North  | Boots   |    1200 |     0
 North  | Jackets |     800 |     0
 North  |         |    2000 |     1
 South  | Boots   |     500 |     0
 South  | Jackets |    1500 |     0
 South  |         |    2000 |     1
        |         |    4000 |     3
```

`ROLLUP (region, product)` は、 `GROUPING SETS ((region, product), (region), ())` の省略形です。列の一覧全体、末尾側を一つずつ省いて残る先頭部分、最後に全行を合計する空集合を指定します。PostgreSQLの文書は、部署別、事業部別、会社全体の給与など、階層的なデータに使う一般的な手段と説明しています。代わりに `CUBE (region, product)` を使うと、すべての部分集合でグループ化するため、全地域を通じた商品ごとの合計も加わります。

ある行のグループ化セットで、グループ化の対象でない列はNULLで返ります。小計の行のセルが空欄なのは、そのためです。 `GROUPING(region, product)` は、それがどの列かを示します。引数ごとに一ビットを割り当て、最も右の引数を最下位ビットとするビットマスクを返し、その行でグループ化していない列のビットを1にします。したがって、0は明細行、1は地域の小計（productをグループ化していない）、3は総計（どちらもグループ化していない）です。レポートは `level` を読み、太字にする行を決めます。

当社のデータ分析・BIダッシュボードサービスは、判断を実際に左右する少数の数値を示し、人が依頼しなくても更新されるダッシュボードとレポートを目指しています。このようなクエリは、そのレポートの下にある層です。合計をデータベースで一度だけ計算し、ダッシュボードは描画だけを行います。

問題が起きるところ

一つの名前で、二つの数値を示します。 売上高、アクティブな顧客、解約率には、それぞれ妥当な定義が複数あり、独自に計算するレポートは、そのうちの一つを選びます。何を集計し、どのデータソースを使い、どの時点の値を示すかを、指標ごとに一度だけ書き留めます。各レポートの数式ではなく、例えばビューとして、データベース内で計算します。Kimballは、この平易な言葉による説明を、ビジネスメタデータと呼びます。どのようなデータがあり、どこから来て、何を意味するかを示すものです。

全員へ、すべてを送ります。 何十ものグラフを載せたダッシュボードと、部署全体へのメーリングリストは、Luhnが止めようとした情報の洪水を再現します。特定の人と、その人の判断から始め、必要な数値を示し、四半期にわたって誰も開かなかったレポートは廃止します。

レポートが、アプリケーションを遅くします。 稼働中のシステムのプライマリーデータベースで重いクエリを動かすと、利用中の顧客と資源を取り合います。レポートは、読み取りレプリカか、定期的にデータを読み込む別のデータウェアハウスで実行します。計算に大きな負担がかかり、頻繁に使う集計は、読み込みのたびに更新するマテリアライズドビューへ置けます。

集計値を、さらに集計します。 各地域の平均注文額を、地域間で平均しても、地域ごとの注文数が同じでなければ、会社全体の平均注文額にはなりません。月ごとの顧客の重複を除いた人数も、四半期分を足せません。七月と八月に購入した人を、二度数えるためです。明細行を保持し、各階層をそこから集計します。グループ化セットは、そのとおりに動きます。上のクエリの各階層は、一つ下の集計結果ではなく、販売の明細行から計算します。

空欄が、二つの意味を持ちます。 地域が不明な販売を一件加えると、上のクエリは、地域が空欄の明細行、その小計として両方のセルが空欄の行、さらに両方が空欄の総計行を返します。`CUBE` では、もっと分かりにくくなります。地域が空欄で商品がBootsの行が二つでき、一つは地域不明の場所で売れたブーツ、もう一つは全地域のブーツを示します。PostgreSQLの文書は、元のデータのNULL

と、グループ化の対象から外した列のNULLを見分けられないと警告しています。小計行の判定には、`IS NULL` ではなく、必ず `GROUPING()` を使います。また、データを読み込む際に、欠損値には「不明」などの見えるラベルを付けます。

昨日の数値を、今日のものとして読みます。 夜間の読み込みが失敗すると、ダッシュボードは最後の正常な数値を表示し続け、古いことを画面上で示しません。どのレポートにも、最後に読み込みが成功した時刻を表示します。また、予定された読み込みが完了しない場合は、担当者を特定して、監視からアラートを送るようにします。

参考資料

- A Business Intelligence System：1958年10月のIBM Journal of Research and Developmentに掲載された、H. P. Luhnの論文です。この用語の出典です。
- Kimball Technical DW/BI System Architecture：バックルーム、提示用の領域、BIアプリケーション、メタデータについての、Kimball Groupの構成説明です。
- テーブル式：`GROUPING SETS` 、 `ROLLUP` 、 `CUBE` と、グループ化していない列をNULLで示す仕組みについての、PostgreSQLのリファレンスです。
- 集約関数：`GROUPING` 関数とそのビットマスクを、例とともに説明します。

用語集

実際の制約に合わせて選ぶ

代表的なデータセット、そこから答えを得たい問い、現在のボトルネックをお知らせください。品質、処理量、復旧の測り方を合意できます。

料金を見る お問い合わせ →

用語を実務に結び付ける

「ビジネスインテリジェンスとは？」がお客様の組織にどう当てはまるかをご相談ください。現在のシステムの具体例が一つあれば、話を始められます。

サービスを見る

welcome@engineer.company +45 50 10 61 84 デンマーク、コペンハーゲン

このページを持ち歩く

連絡先カード

心を込めて丁寧に制作： D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

著作権 © 2025 – 現在 · Engineer ApS · すべての権利を留保します。

印刷元： https://engineer.company/ja/glossary/business-intelligence/